Safety in Integrated Systems Health Engineering and Management

Nancy G. Leveson

MIT

- You’ve carefully thought out all the angles.
- You’ve done it a thousand times.
- It comes naturally to you.
- You know what you’re doing, it’s what you’ve been trained to do your whole life.
- Nothing could possibly go wrong, right?

Think Again.

Reliability Engineering vs. System Safety

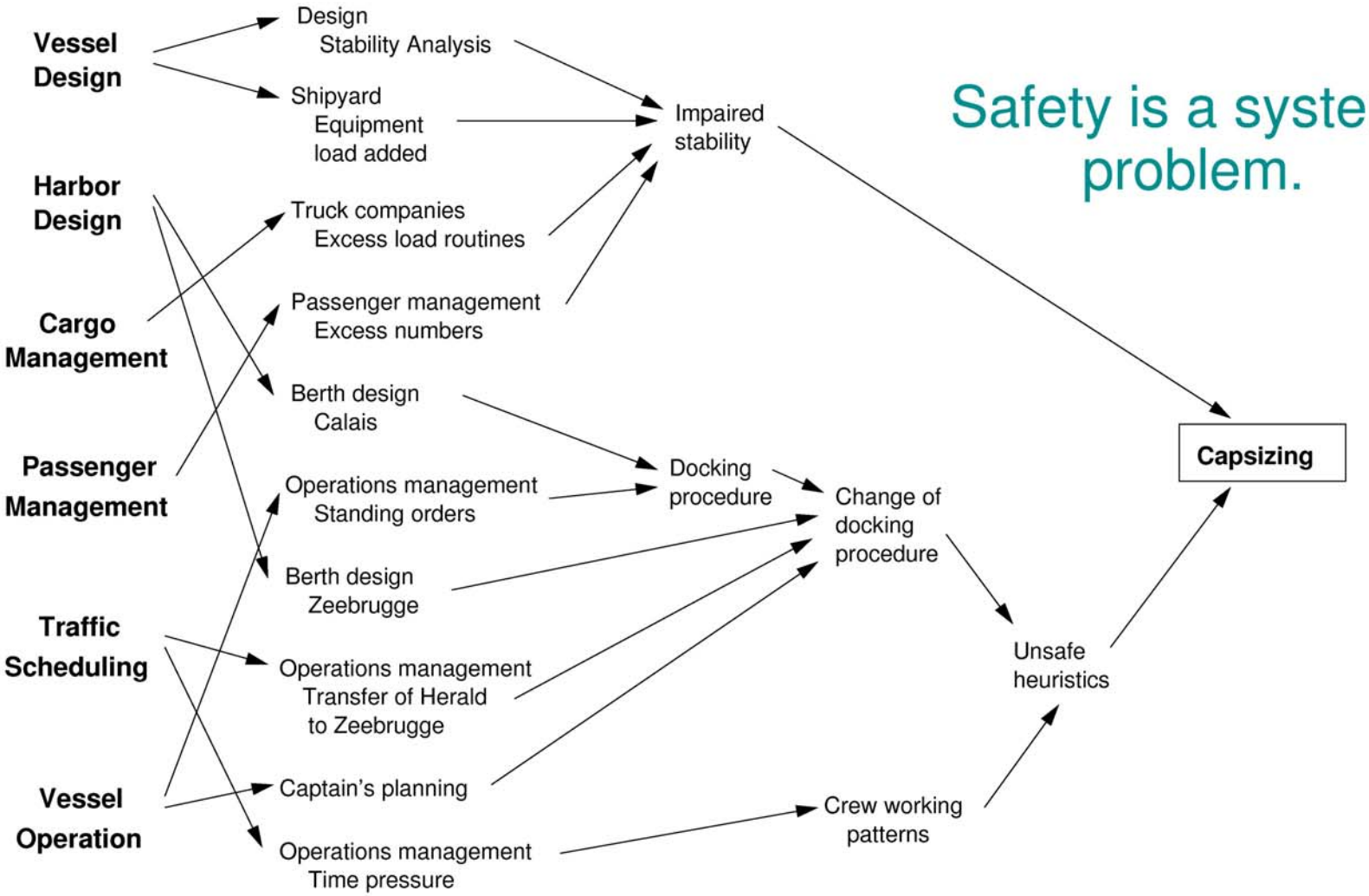
- Both arose after World War II
- Reliability engineering often confused with system safety engineering, but they are different and sometimes even conflict
- Reliability engineering focuses on preventing failure
- System safety focuses on eliminating and controlling hazards
  - Considers interactions among components and not just component failures
  - Includes non-technical aspects of systems
- Highly reliable systems may be unsafe and safe systems may not be reliable.

Traditional Chain-of-Events Accident Causality Models

- Explain accidents in terms of multiple events, sequenced as a forward chain over time.
- Events linked together by direct relationships (ignore indirect, non-linear relationships).
- Events almost always involve component failure, human error, or energy-related events.
- Form the basis for most safety-engineering and reliability engineering analysis (FTA, FMEA, PRA) and design.

Limitations of Event-Chain Causality Models

- Social and organizational factors
- System accidents
- Software Error
- Human Error
  - Cannot effectively model human behavior by decomposing it into individual decisions and actions and studying it in isolation from
    - physical and social context
    - value system in which it takes place
    - dynamic work process
- Adaptation
  - Major accidents involve systematic migration of organizational behavior to higher levels of risk.



Safety is a system problem.

Operational Decision Making:

Decision makers from separate departments in operational context very likely will not see the forest for the trees.

Accident Analysis:

Combinatorial structure of possible accidents can easily be identified.

New Approaches Based on Systems Theory

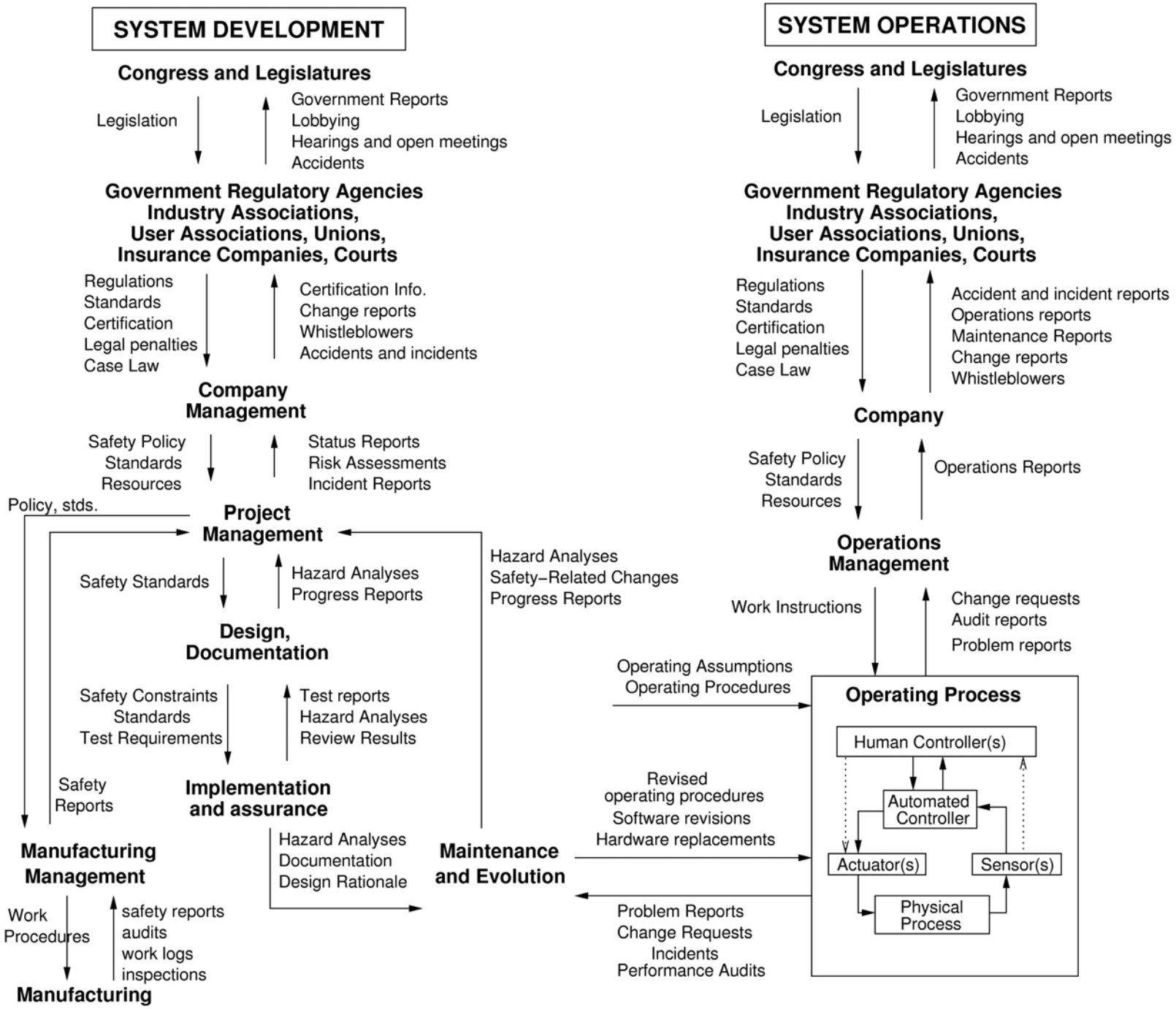
- Rasmussen – Hierarchical model of accident causation
- STAMP (Systems Theoretic Accident Modeling and Processes
  - New accident causation model based on systems theory
  - New hazard analysis technique (STPA)
    - Works for hardware, software, human error, social factors, management errors, etc.
    - Includes what we do now, but more
  - New, more powerful risk management tools (including policy analysis and evaluation and “canary in the coal mine”)
  - Designing for safety
  - Root cause analysis and incident/accident investigation

A Systems Theory Model of Accidents

- Accidents arise from interactions among humans, machines, and the environment.
  - Not simply chains of events or linear causality, but more complex types of causal connections.
- Safety is an emergent property that arises when components of system interact with each other within a larger environment.
  - A set of constraints related to behavior of components in system enforces that property.
  - Accidents when interactions violate those constraints (a lack of appropriate constraints on the interactions).
  - Software as a controller embodies or enforces those constraints.

A Systems Theory Model of Accidents

- Systems should not be treated as a static design
  - A socio–technical system is a dynamic process continually adapting to achieve its ends and to react to changes in itself and its environment
  - Preventing accidents requires designing a control structure to enforce constraints on system behavior and adaptation.


SYSTEM DEVELOPMENT
Congress and Legislatures
Legislation
Government Reports
Lobbying
Hearings and open meetings
Accidents
Government Regulatory Agencies
Industry Associations,
User Associations, Unions,
Insurance Companies, Courts
Regulations
Standards
Certification
Legal penalties
Case Law
Certification Info.
Change reports
Whistleblowers
Accidents and incidents
Company
Management
Safety Policy
Standards
Resources
Status Reports
Risk Assessments
Incident Reports
Policy, stds.
Project
Management
Safety Standards
Hazard Analyses
Progress Reports
Design,
Documentation
Safety Constraints
Standards
Test Requirements
Test reports
Hazard Analyses
Review Results
Safety
Reports
Implementation
and assurance
Hazard Analyses
Documentation
Design Rationale
Manufacturing
Management
Work
Procedures
safety reports
audits
work logs
inspections
Manufacturing
Hazard Analyses
Safety−Related Changes
Progress Reports
Maintenance
and Evolution
SYSTEM OPERATIONS
Congress and Legislatures
Legislation
Government Reports
Lobbying
Hearings and open meetings
Accidents
Government Regulatory Agencies
Industry Associations,
User Associations, Unions,
Insurance Companies, Courts
Regulations
Standards
Certification
Legal penalties
Case Law
Accident and incident reports
Operations reports
Maintenance Reports
Change reports
Whistleblowers
Company
Safety Policy
Standards
Resources
Operations Reports
Operations
Management
Work Instructions
Change requests
Audit reports
Problem reports
Operating Assumptions
Operating Procedures
Operating Process
Human Controller(s)
Automated
Controller
Actuator(s)
Sensor(s)
Physical
Process
Revised
operating procedures
Software revisions
Hardware replacements
Problem Reports
Change Requests
Incidents
Performance Audits

A Systems Theory Model of Accidents (3)

- Views accidents as a control problem

  e.g., O–ring did not control propellant gas release by sealing gap in field joint

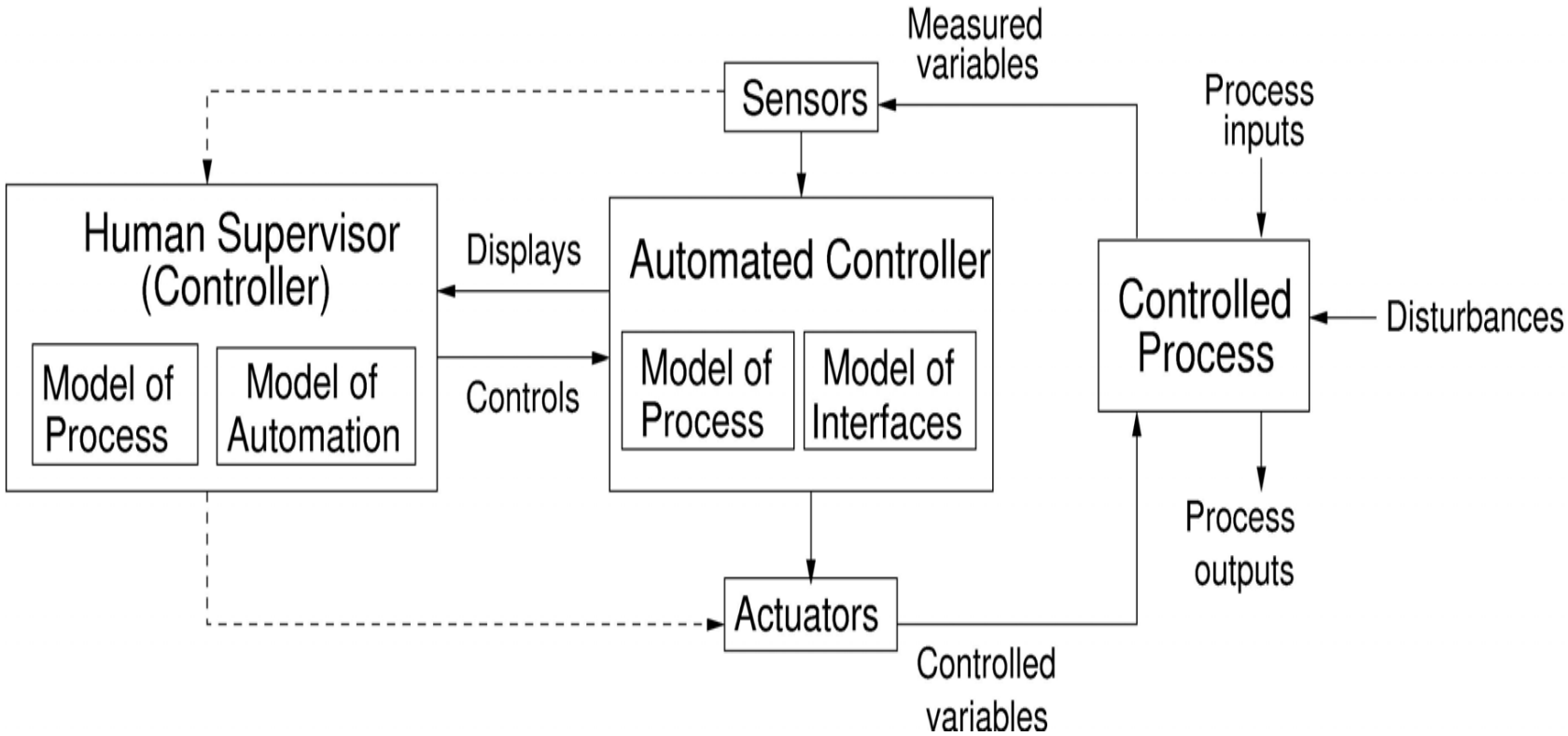
  Software did not adequately control descent speed of Mars Polar Lander.

- Events are the result of the inadequate control

  Result from lack of enforcement of safety constraints

- To understand accidents, need to examine control structure itself to determine why inadequate to maintain safety constraints and why events occurred.

Not a "blame" model – trying to understand "why"



Process models must contain:

- Required relationship among process variables
- Current state (values of process variables)
- The ways the process can change state

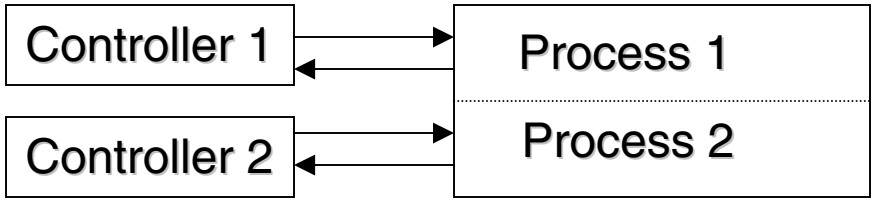
Some Causal Factors in Accidents

- Design does not enforce safety constraints.
  - mishandled disturbances, failures, dysfunctional interactions
- Controller provides inadequate control actions
  - inconsistent/incorrect process models
  - inadequate or missing feedback
  - inadequate control algorithms
  - time lags

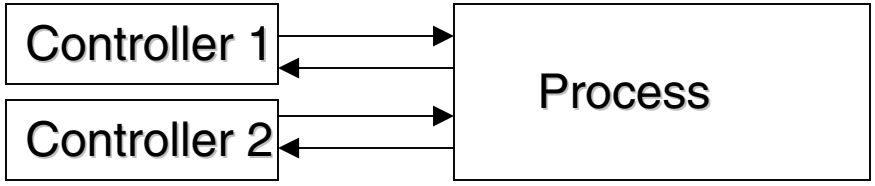
Some Causal Factors in Accidents (2)

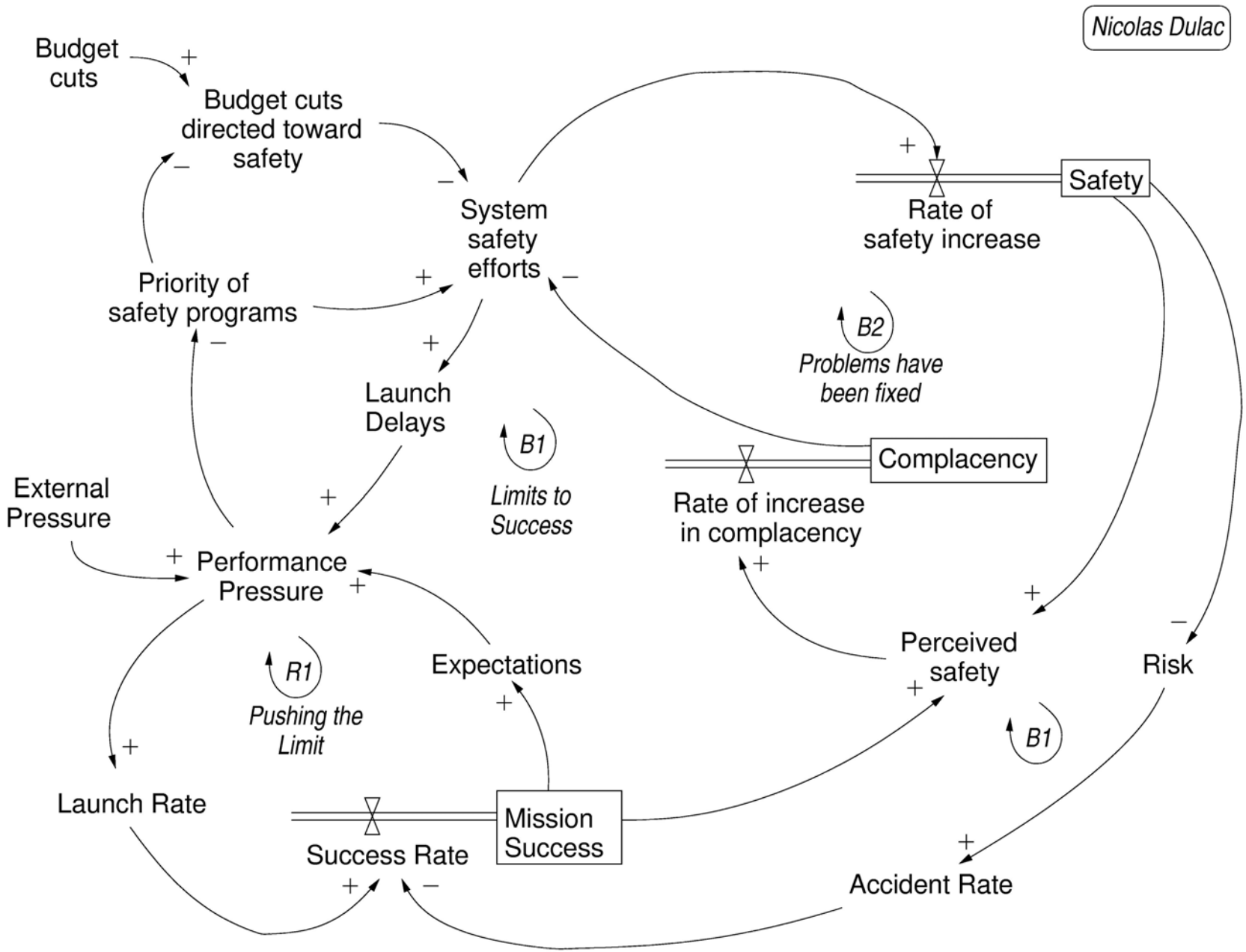
- Control structure degrades over time, asynchronous evolution.

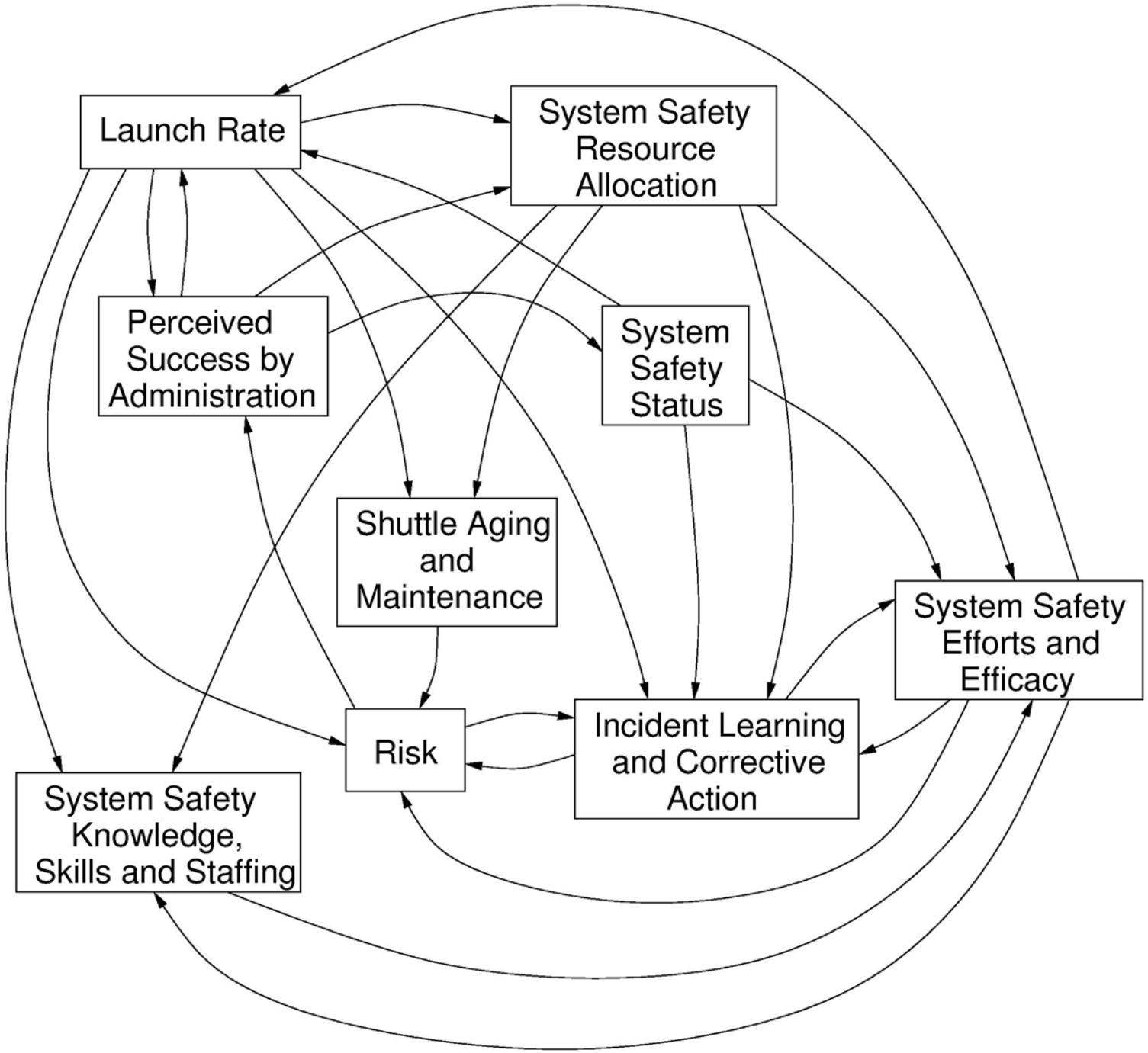
- Control actions inadequately coordinated among multiple controllers.
  - Boundary areas



  - Overlap areas (side effects of decisions and control actions)



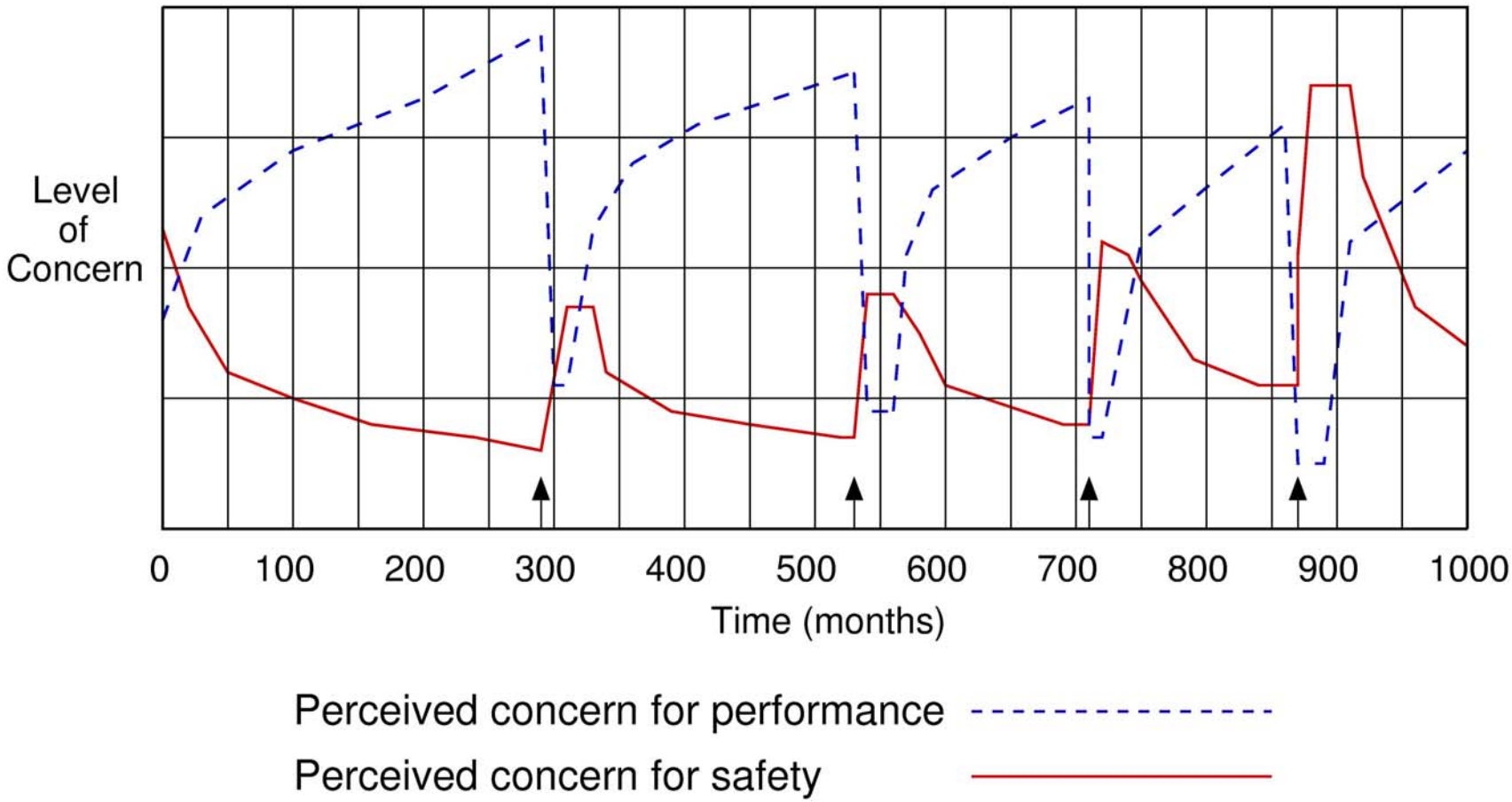

Nicolas Dulac
Budget cuts
Budget cuts directed toward safety
System safety efforts
Rate of safety increase
Safety
Priority of safety programs
Launch Delays
B2
Problems have been fixed
B1
Limits to Success
Complacency
Rate of increase in complacency
External Pressure
Performance Pressure
Perceived safety
Risk
R1
Pushing the Limit
Expectations
B1
Launch Rate
Mission Success
Success Rate
Accident Rate


Launch Rate
System Safety Resource Allocation
Perceived Success by Administration
System Safety Status
Shuttle Aging and Maintenance
System Safety Efforts and Efficacy
Risk
Incident Learning and Corrective Action
System Safety Knowledge, Skills and Staffing

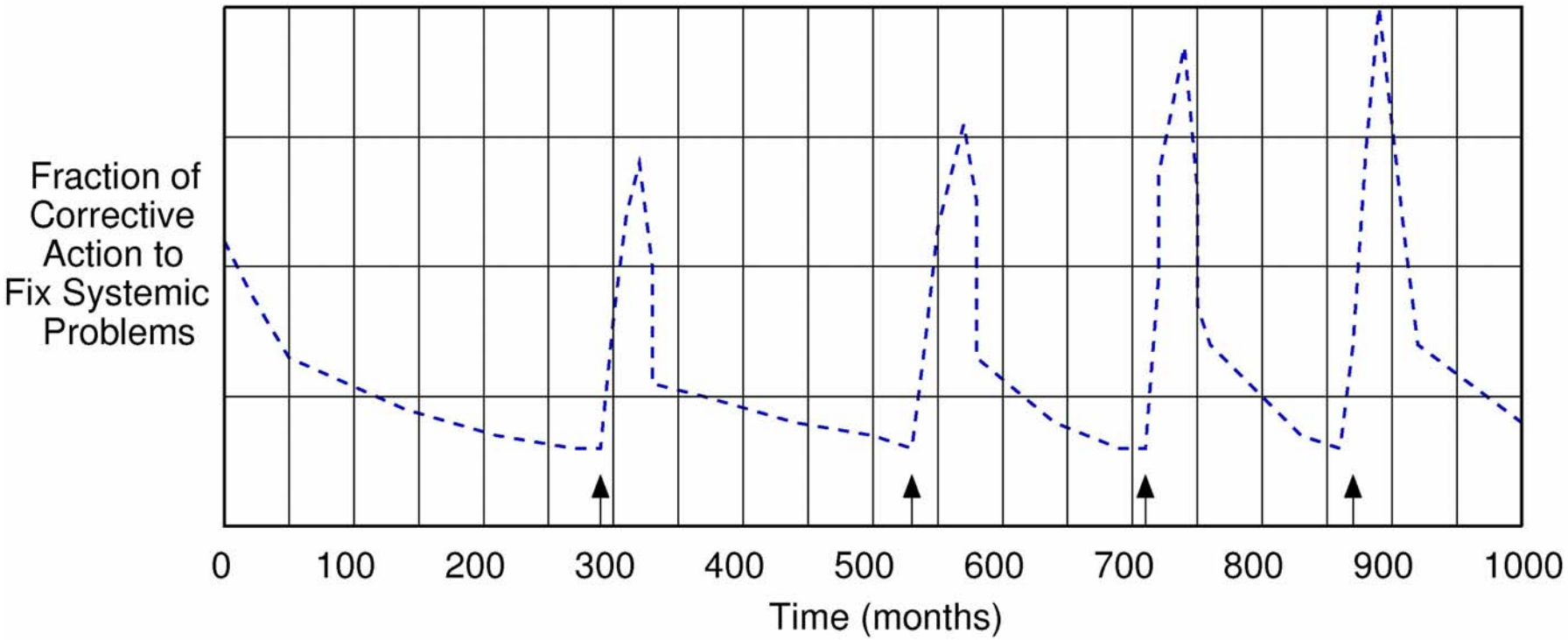
How can this model help us?

- It allows us to
  - Understand how and why accidents have occurred
  - Test and validate changes and new policies
  - Learn which “levers” have a significant and sustainable effect
  - Facilitate the identification and tracking of metrics to detect increasing risk

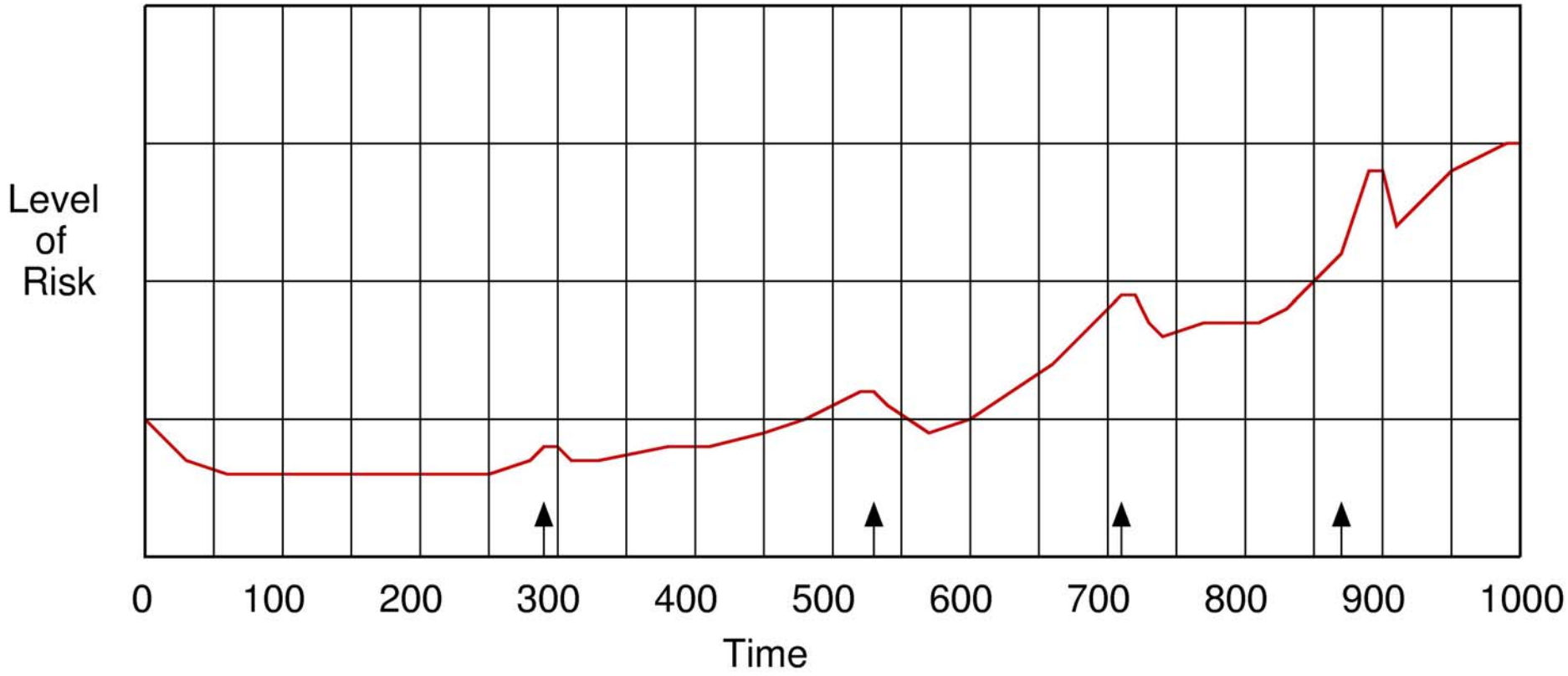
Accidents lead to a re–evaluation of NASA safety and performance priorities but only for a short time:



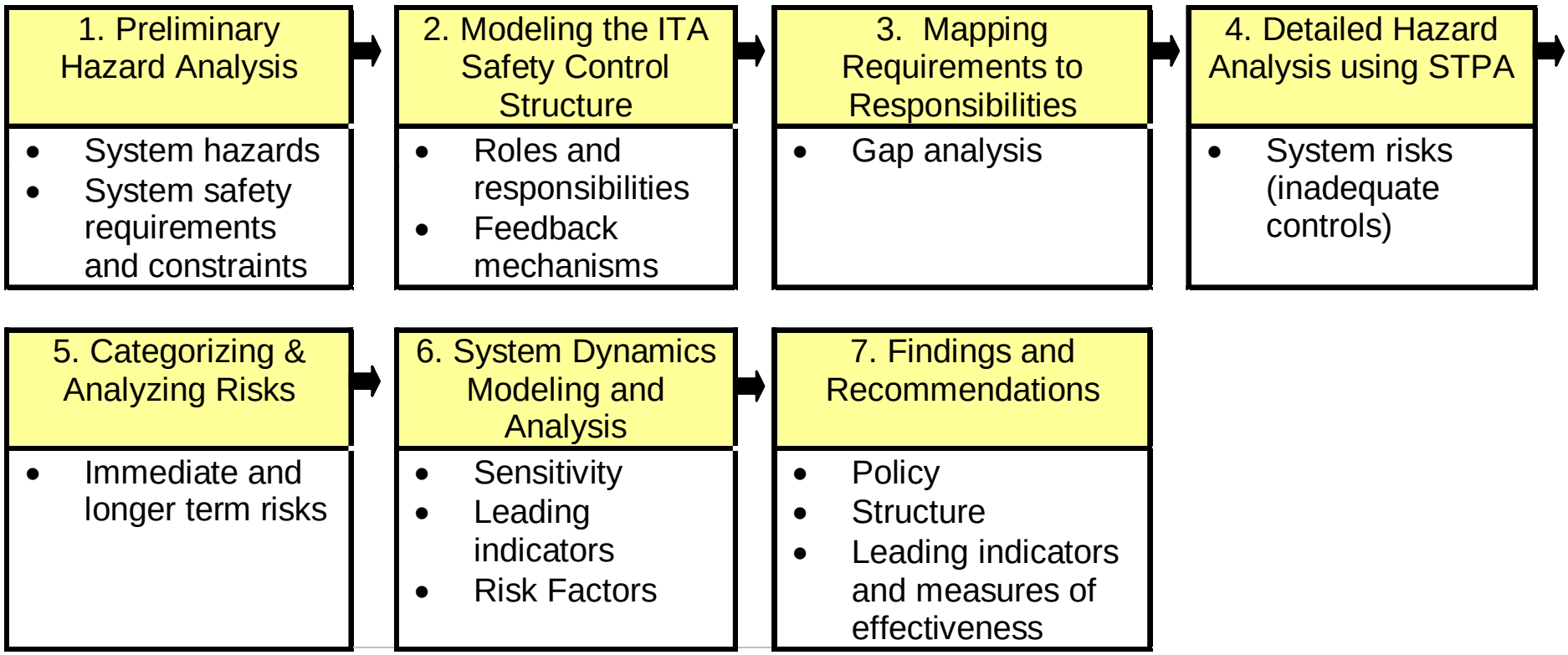
Attention to fixing systemic problems lasts only a short time after an accident



Responses to accidents have little lasting impact on risk



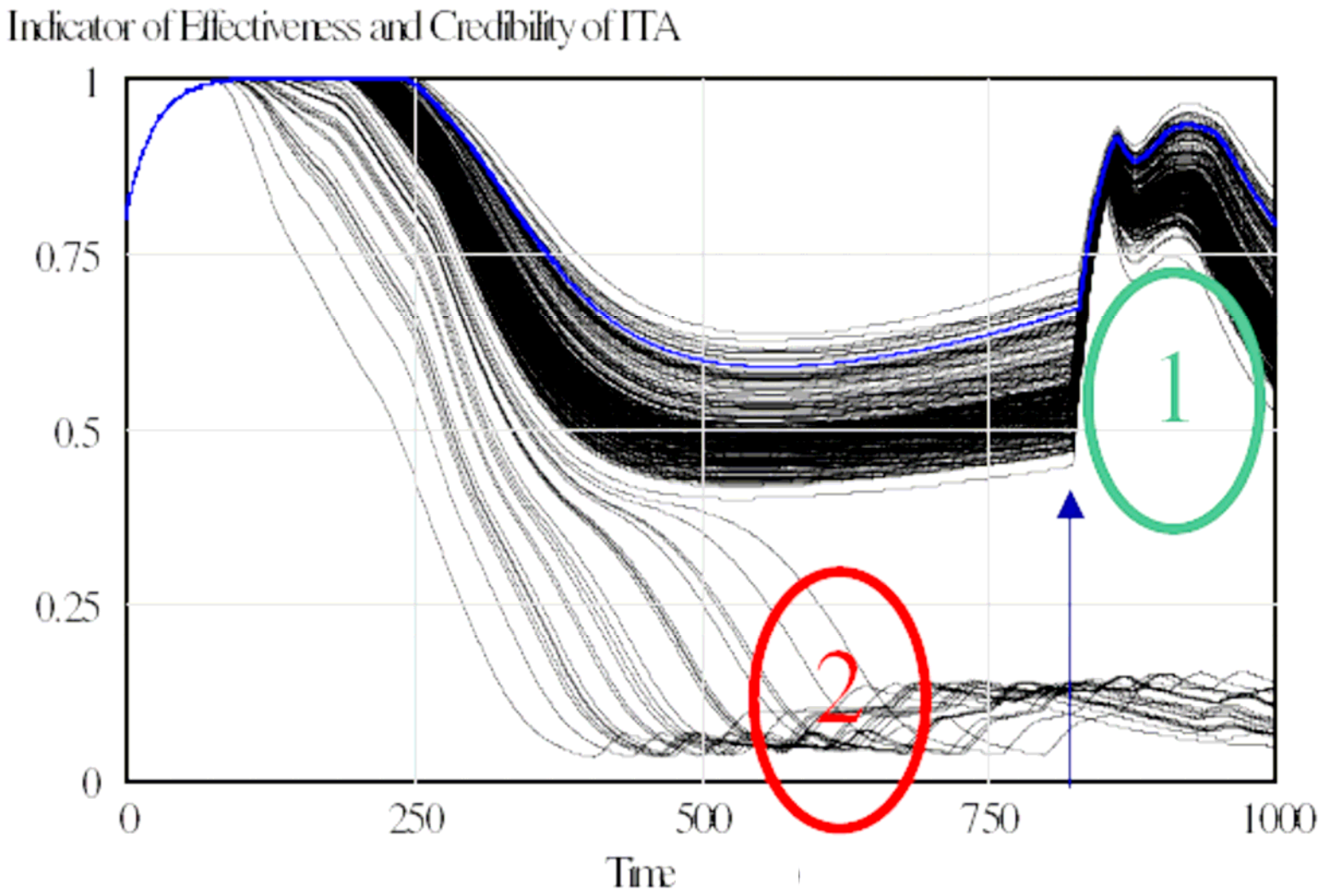
The Process

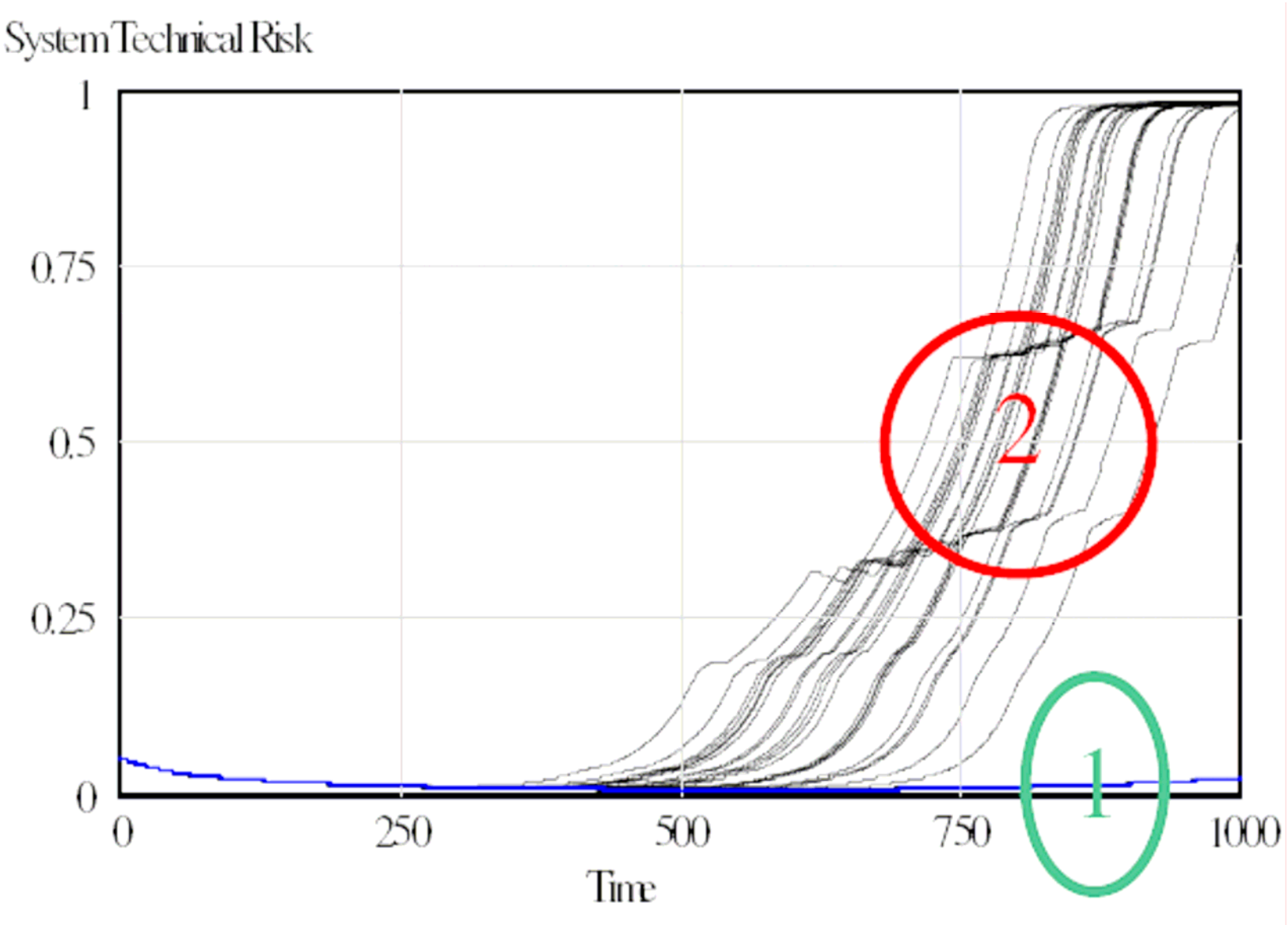

1. Preliminary Hazard Analysis
• System hazards
• System safety requirements and constraints
2. Modeling the ITA Safety Control Structure
• Roles and responsibilities
• Feedback mechanisms
3. Mapping Requirements to Responsibilities
• Gap analysis
4. Detailed Hazard Analysis using STPA
• System risks (inadequate controls)
5. Categorizing & Analyzing Risks
• Immediate and longer term risks
6. System Dynamics Modeling and Analysis
• Sensitivity
• Leading indicators
• Risk Factors
7. Findings and Recommendations
• Policy
• Structure
• Leading indicators and measures of effectiveness

Example Result

- ITA has potential to significantly reduce risk and to sustain an acceptable risk level
- But also found significant risk of unsuccessful implementation of ITA that needs to be monitored
  - 200-run Monte-Carlo sensitivity analysis
  - Random variations of +/- 30% of baseline exogenous parameter values

Sensitivity Analysis Results




System Technical Risk
1
0.75
0.5
0.25
0
0
250
500
750
1000
Time
2
1

Successful Scenarios

- Self-sustaining for short period of time if conditions in place for early acceptance.
- Provides foundation for a solid, sustainable ITA program implementation under right conditions.
- Successful scenarios:
  - After period of high success, effectiveness slowly declines
    - Complacency
    - Safety seen as solved problem
    - Resources allocated to more urgent matters
  - But risk still at acceptable levels and extended period of nearly steady-state equilibrium with risk at low levels

Unsuccessful Implementation Scenarios

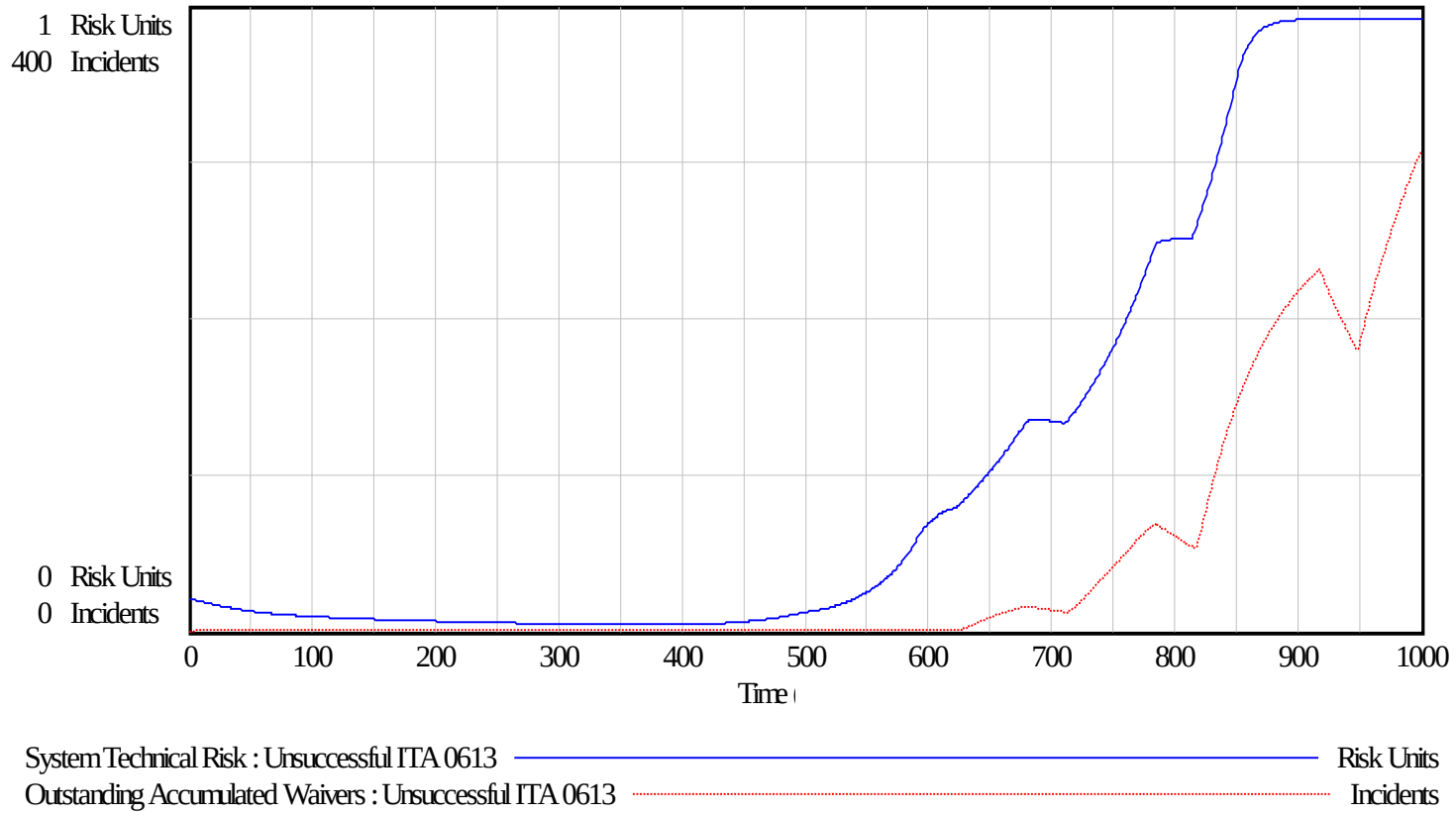
- Effectiveness quickly starts to decline and reaches unacceptable levels
  - Limited ability of ITA to have sustained effect on system
  - Hazardous events start to occur, safety increasingly perceived as urgent problem
  - More resources allocated to safety but TA and TWHs have lost so much credibility they cannot effectively contribute to risk mitigation anymore.
  - Risk increases dramatically
  - ITA and safety staff overwhelmed with safety problems
  - Start to approve an increasing number of waivers so can continue to fly.

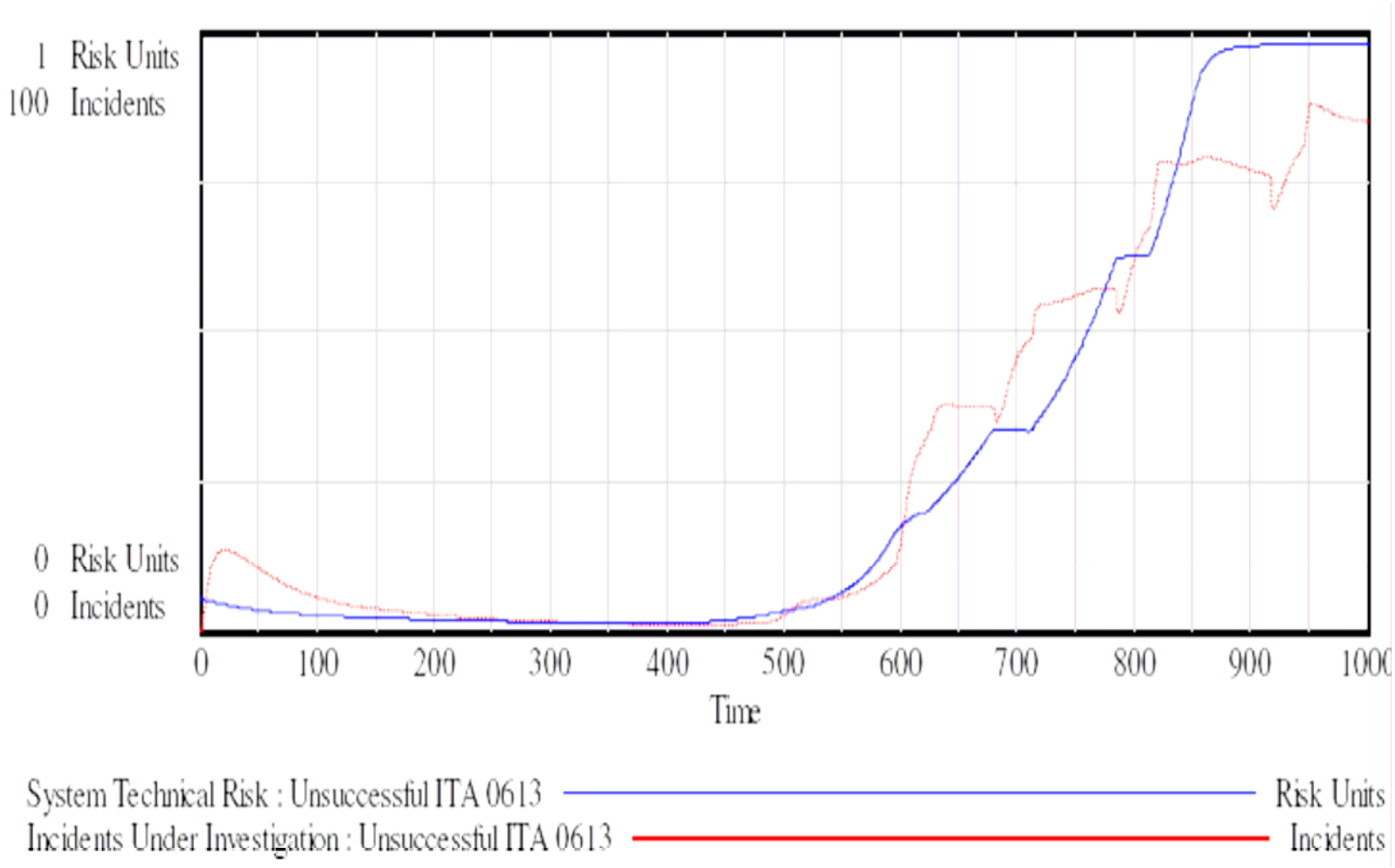
Unsuccessful Scenario Factors

- As effectiveness of ITA decreases, number of problems increase
  - Investigation requirements increase
  - Corners may be cut to compensate
    - Results in lower-quality investigation resolutions and corrective actions
  - TWHs and Trusted Agents become saturated and cannot attend to each investigation in timely manner
  - Bottleneck created by requiring TWHs to authorize all safety-related decisions, making things worse
- Want to detect this reinforcing loop while interventions still possible and not overly costly (resources, downtime)

Lagging vs. Leading Indicators

- Number of waivers issued good indicator but lags rapid increase in risk




1 Risk Units
100 Incidents
0 Risk Units
0 Incidents
0
100
200
300
400
500
600
700
800
900
1000
Time
System Technical Risk : Unsuccessful ITA 0613
Risk Units
Incidents Under Investigation : Unsuccessful ITA 0613
Incidents